Australian Government
Department of Defence
Defence Science and
Technology Organisation

An Investigation of the Pareto Distribution as a Model for High Grazing Angle Clutter

Graham V. Weinberg

Electronic Warfare and Radar Division

Defence Science and Technology Organisation

DSTO–TR–2525

ABSTRACT

The Pareto Distribution is proposed as a new model for high resolution, high grazing angle radar sea clutter returns. It is assessed using the Ingara radar clutter gathered in 2004. Included is range profile modelling as well as global distributional fitting to clutter streams.

APPROVED FOR PUBLIC RELEASE

Published by

DSTO Defence Science and Technology Organisation

PO Box 1500

Edinburgh, South Australia 5111, Australia

Telephone: (08) 7389 5555

Facsimile: (08) 7389 6567

AR No. AR 014-947

March 2011

APPROVED FOR PUBLIC RELEASE

An Investigation of the Pareto Distribution as a Model for High Grazing Angle Clutter

Executive Summary

Clutter models are a vital tool in the validation of the performance of radar detection schemes under controlled conditions. Complicated clutter models result in mathematical difficulties in the determination of optimal and even suboptimal radar detection schemes. Hence it is desirable to see whether this complexity can be minimised through the selection of simpler clutter models. Recently, it was proposed that the well-known Pareto Distribution is an acceptable model for high resolution maritime radar clutter, at low grazing angles. This was validated through trials data, and it was found that the Pareto Distribution perfomed comparably to other modern models of sea clutter returns, namely the K- and KK-Distributional models.

Using the Defence Science and Technology Organisation's Intelligence, Surveillance and Reconnaissance Division's Ingara Data, which is X-Band fully polarmetric high resolution, high grazing angle radar clutter, it is shown that the Pareto Distribution is also an appropriate model. The importance of this discovery is that in future detection work in support of Project AIR 7000, a simpler clutter model can be employed, facilitating the construction of optimal and suboptimal detectors for scenarios of interest to Task 07/040.

Author

Graham V. Weinberg

EWRD

Graham V. Weinberg is an S&T5 Mathematician working in the Microwave Radar Branch of EWRD. He holds a PhD in applied probability and stochastic processes from the University of Melbourne (2001) and an MSc in Defence Signal and Information Processing from the University of Adelaide (2010).

Contents

Figures

Tables

1 The Pareto Distribution

1.1 Background

The Pareto Distribution [1-3] is named after the Italian economist Vilfredo Pareto (15 July 1848 – 19 August 1923) [4, 5], and is a power law probability distribution that has been found to be an excellent model of long tailed phenomena [6]. Its original application was in the modelling of income over a population [7]. It has been used in the modelling of actuarial data; an example is in excess of loss quotations in insurance [8]. Its usefulness as a model for long tailed distributions has resulted in applications to a number of diverse areas including physics, hydrology and seismology [6]. It has also found suitable applications in engineering, such as internet teletraffic modelling. Specifically, it has found application to the characterisation of file sizes, TCP packet inter-arrival times, FTP transfer times and burst sizes [9]. The distribution has also found application in sonar [10] and radar [11-13] as a model of sea clutter returns.

In terms of applications in radar, the Pareto distribution has been used to approximate the tail of a density function for determination of the likelihood ratio test for radar detection [11]. In the context of modelling spiky sea clutter returns, [12] used the generalised Pareto Distribution to model clutter tails and designed a Constant False Alarm Rate (CFAR) detector on the basis of this. More recently, [13] examined the application of the generalised Pareto Distribution to the modelling of high resolution X-band radar sea clutter, obtained at a low grazing angle. This work showed that a two-parameter Pareto distribution fitted polarised clutter returns quite well. This distributional fit was also compared to a number of classical radar clutter amplitude models (Lognormal, Weibull) as well as more contemporary ones (K, KK and WW, where the latter two are mixture distributions). It is demonstrated in [13] that the Pareto Distribution provides a better fit to their clutter data than the traditional models and the K and WW-Distributions. They also reported its performance is very close to that of the KK-Distribution. Since the Pareto Distribution is characterised by a simple two parameter density function, these results are quite promising, as remarked in [13]. To clarify why, recall the KK-Distribution is a 4 to 5 parameter family with a complicated distribution function. The latter involves

modified Bessel functions, making it difficult to employ in radar detection schemes. The Pareto Distribution is amenable to mathematical analysis, and hence presents itself as an attractive alternative to the KK-Distribution.

1.2 The Clutter Set

The sea clutter used for this report was collected by DSTO's Intelligence, Surveillance and Reconnaissance Division (ISRD) using their radar testbed called Ingara. This is an X-Band fully polarised radar. The 2004 trial used to collect the data was located in the Southern Ocean, roughly 100km south of Port Lincoln in South Australia. Details of this trial, as well as data analysis of the sea clutter, can be found in [14-16].

The radar operated in a circular spotlight mode, so that the same patch of sea surface was viewed at different elevation angles, as well as different azimuth angles. The radar used a centre frequency of 10.1 GHz, with 20 μs pulse width. Additionally, the radar operated at an altitude of 2314m for a nominal incidence angle of 50°, and at 1353m for 70° incidence angle. The trial collected data at incidence angles varying from 40° to 80°, on 8 different days over an 18 day period. As in [16], we focus on data from two particular flight test runs. These correspond to run34683 and run34690, which were collected on 16 August 2004 between 10:52am and 11:27am local time [16]. Dataset run34683 was obtained at an incidence angle of 51.5°, while run34690 was at 67.2°. In terms of grazing angles, these correspond to 38.7° and 22.8° respectively. Each of these datasets were also processed to cover azimuth angle spans of 5° over the full 360° range. Roughly 900 pulses were used, and 1024 range compressed samples for each pulse were produced, at a range resolution of 0.75m.

In [16] parameter estimates for the data sets run34683 and run34690 are given, enabling the fitting of the K- and KK-Distributions to this data. This will be employed in the numerical analysis to follow.

1.3 Research Objectives

The primary objective of this research is to investigate whether the same results, as recorded in [13], apply for the case of X-Band High Resolution maritime radar sea clutter returns obtained at a *high grazing angle*. Hence, the statistical objective is to assess how well the Pareto Distribution fits such clutter, and in particular, the Ingara 2004 Trial data. If the Pareto Distribution is a suitable model for this clutter, the construction of optimal and suboptimal detectors for targets in high resolution high grazing angle sea clutter would be dramatically simplified.

Section 2 introduces statistical properties of the Pareto Distribution, as well as methods used for estimation of its parameters. Section 3 analyses the result of distributional fits to the data, including comparative results for the K- and KK-Distributions. Included are analyses of range profile fits as well as fits to an entire ensemble of data.

A useful reference on parameter estimation is [2], which includes both frequentist as well as Bayesian estimation details specialised to the Pareto class of distributions. Additionally, [8] has an excellent account of maximum likelihood parameter estimation for the Pareto Distribution.

2 Pareto Distributional Analysis

Before analysing the Ingara data fit, the Pareto Distribution is introduced statistically, and a number of its characteristics are given. Additionally, methods of estimating its shape and scale parameters are discussed.

2.1 Statistical Definition and Properties

The mathematical properties of the Pareto Distribution are now given. The reader can consult [2] and [3] for further details if required. Let α and β be positive constants. Then a random variable X has a Pareto Distribution, written $X \stackrel{d}{=} Pa(\alpha, \beta)$ if its density is

$$f_X(t) = \begin{cases} 0 & t < \beta, \\ \frac{\alpha\beta^\alpha}{t^{\alpha+1}} & t \geq \beta. \end{cases} \tag{2.1}$$

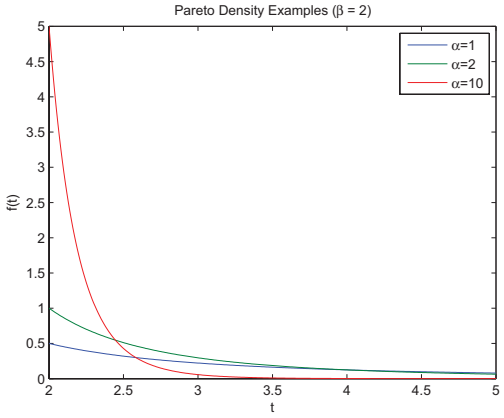

Figure 2.1: *Examples of the Pareto Distribution's density. All cases use a scale parameter of* $\beta = 2$.

The parameter α is known as the distribution's shape parameter, while β is called the scale parameter. The latter specifies where on the real line the distribution's support begins, which is the interval $[\beta, \infty)$. The shape parameter controls how fast the distribution's tail decreases. Figure 2.1 illustrates this graphically, for the case where the support is $[2, \infty)$ and for $\alpha \in \{1, 2, 10\}$.

Its cumulative distribution function can be derived by integrating (2.1); in particular, it is given by

$$F_X(t) = \begin{cases} 0 & t < \beta \\ 1 - \left(\frac{\beta}{t}\right)^{\alpha} & t \geq \beta. \end{cases} \tag{2.2}$$

The generalised Pareto Distribution is often examined in the literature, and so we briefly examine it here, adopting the formulation in [13]. This distribution has probability density

$$f(t) = \frac{1}{\lambda}\left(1 - \frac{\omega t}{\lambda}\right)^{\frac{1}{\omega}-1}, \tag{2.3}$$

where ω is called the shape parameter and λ is the scale parameter. In [13], it is pointed out that this distribution has support $0 \leq t < \infty$ when the shape parameter is non-positive, and $0 \leq t \leq \frac{\lambda}{\omega}$ when ω is positive. In the special case where $\omega = 0$, the density can be shown to limit to an Exponential Distribution's density with mean λ. Additionally, when $\omega = 1$ it reduces to a uniform density on $[0, \lambda]$. As in [13], the case where $\omega < 0$ is

considered only, and through a very simple change of variables[1] it can be shown that (2.3) is equivalent to (2.1). Matlab^{TM} returns parameter estimates for the Pareto Distribution with density (2.3), and so in the distributional fitting to follow, parameters cited will be for the density (2.3).

The existence of moments of the Pareto Distribution depends on the parameter α. In particular, with reference to the density (2.1), the mean is

$$\mathbb{E}(X) = \frac{\alpha\beta}{\alpha - 1} \tag{2.4}$$

for $\alpha > 1$, while the variance

$$\mathbf{var}(X) = \frac{\beta^2\alpha}{(\alpha-1)^2(\alpha-2)} \tag{2.5}$$

exists for $\alpha > 2$.

The Pareto Distribution is intimately related to the Exponential Distribution. If we define a random variable $Y = \log\left(\frac{X}{\beta}\right)$, then it is not difficult to show by constructing the distribution function of Y that it has an Exponential Distribution with parameter α. This means that a Pareto $Pa(\alpha, \beta)$ random variable is statistically equivalent to βe^Y, where $Y \stackrel{d}{=} Exp(\alpha)$.

The Pareto Distribution requires the estimation of two parameters in order to fit it to a data set. This is examined in the next Subsection.

2.2 Parameter Estimation

The Pareto distribution has a power law density, and so exhibits a long tail in its distribution. Due to this fact, moment-based estimates, such as the mean and variance, are likely to be affected. Hence, method of moments estimators are not considered appropriate for this distributional fitting. However, maximum likelihood estimators (MLEs) for such a distribution have useful properties, which will be demonstrated in the following. We begin by constructing the MLE for the scale parameter.

Suppose we have a random sample of observations $\{x_1, x_2, \ldots, x_n\}$, obtained from a common Pareto Distribution. We denote the corresponding joint random variables as

[1] Define $\alpha = -\frac{1}{\omega} > 0$ and change variables to $z = t + \lambda\alpha$.

$\{X_1, X_2, \ldots, X_n\}$. Then by analysing the likelihood function of a series of independent and identically distributed Pareto random variables, the maximum likelihood estimator of the scale parameter can be shown to be

$$\hat{\beta} = x_{(1)} = \min\{x_1, x_2, \ldots, x_n\}. \tag{2.6}$$

This result is not surprising, since it is plausible that the finite endpoint of the distribution's support should be estimated from the smallest observation in a random sample from the distribution.

It can also be demonstrated that the conditional distribution of the random sample $X_1, X_2, \ldots, X_n$, conditioned on the event that the random variable $X_{(1)} = t$, does not depend on β, implying that the mimimum $X_{(1)}$ is a sufficient statistic for β. Furthermore, the Pareto Distribution belongs to the exponential family of distributions, which is complete. This implies that we can construct the minimum variance unbiased estimator. By observing that $\mathbb{E}(X_{(1)}) = \frac{n\alpha\beta}{n\alpha-1}$, we can then choose the random variable

$$\widehat{B} = \frac{(n\alpha - 1)X_{(1)}}{n\alpha}. \tag{2.7}$$

Since this is a function of a complete sufficient statistic, it is a minimum variance unbiased estimator for its mean, which is clearly β by construction. For large sample sizes, the statistic $\widehat{B}$ is asymptotically the same as the minimum $X_{(1)}$, and so this can be used to remove dependence on α in the estimator (2.7).

Next, we examine the MLE for the shape parameter. By analysing the likelihood function, the MLE of the shape parameter α can be shown to be

$$\hat{\alpha} = \frac{n}{\sum_{j=1}^{n}\left[\log(x_j) - \log(\widehat{B})\right]}. \tag{2.8}$$

Due to the fact that $\log(X_j)/\widehat{B}$ is Exponentially Distributed (as remarked previously), it can be shown that $T = \sum_{j=1}^{n}\left[\log(X_j) - \log(\widehat{B})\right]$ has a Gamma Distribution with density

$$f_T(t) = \frac{\alpha^n}{(n-1)!}t^{n-1}e^{-\alpha t}, \tag{2.9}$$

and consequently it can be shown, using the properties of Gamma Distribution moments [2], that the estimator $\widehat{\Theta} = \frac{n}{T}$ of α has moments

$$\mathbb{E}(\widehat{\Theta}) = \frac{n}{n-1}\alpha. \tag{2.10}$$

Hence the adjusted estimator

$$\widehat{A} = \frac{n-1}{T} = \frac{n-1}{\sum_{j=1}^{n}\left[\log(x_j) - \log(\widehat{B})\right]}, \tag{2.11}$$

is unbiased for α. By constructing joint conditional distributions, it can be demonstrated that $\widehat{A}$ is a complete sufficient statistic for α, and since it is unbiased, it is also the minimum variance unbiased estimator for α (also due to completeness as before).

Hence (2.7) can be used to estimate the scale parameter (using a large sample approximation), while (2.11) is used with this to estimate the shape parameter.

3 Distributional Fitting

An analysis is now undertaken of the suitability of the Pareto model for High Grazing Angle sea clutter. The data sets used for this purpose, as remarked previously, correspond to run34683 and run34690. Specifically, the subsets designated run34683_rccal_190, run34683_rccal_225, run34690_rccal_005 and run34690_rccal_255 are the focus. In each of these data sets, the last numerical label refers to the azimuth angle of the radar at the data capture instant. The analysis of data is undertaken at the range profile stage first, followed by fitting to an entire data stream.

3.1 Preliminary Testing: Goodness of Fit Analysis

As a preliminary examination of the suitability of the Pareto model for the Ingara data, a series of goodness of fit tests were performed on the data sets indicated above.. The first series of tests performed were to assess whether range profiles could be modelled by a Pareto Distribution. The range profiles exhibited correlations in time; Figure 3.1 (left plot) shows an example of this, for the HH channel. This is from a range profile generated from the data set run34683_rccal_225. It was decided that sequential decorrelation could be perfomed so that a Kolmogorov-Smirnoff Goodness of Fit test [17] could be run over the entire ensemble of data being analysed. The decorrelation was done by randomising the samples of clutter using a randomisation of their index parameters. Figure 3.1 (right plot)

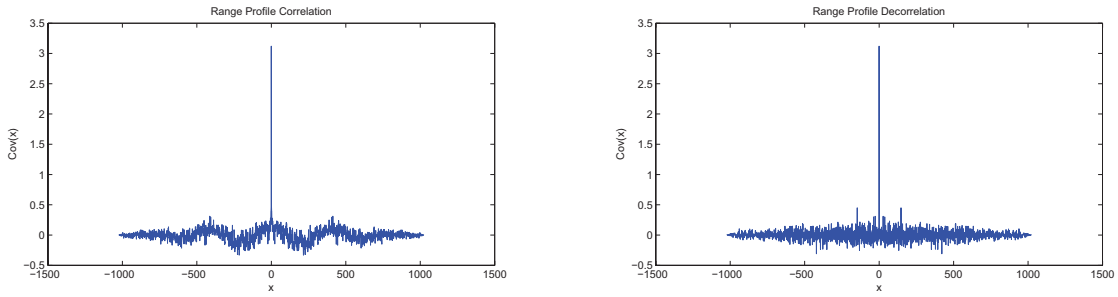

Figure 3.1: *Range profile decorrelation: left plot shows an example of HH-Channel autocovariance, and right plot shows the same result after the data has been decorrelated.*

shows the result of this to the data used to produce the left range profile's autocovariance function.

It is important to note that the data also exhibits correlations between range profiles, since these will correspond to returns to the radar from pulses sent at slightly different azimuth angles, and so will reflect off overlapping patches of sea surface. Hence we do not attempt a goodness of fit test applied to an entire data stream.

The first test was perfomed on the data set run34690_rccal_255, which contained 966 pulses for each range compressed sample. The result of the Kolmogorov-Smirnoff test is shown in Table 3.1, which records P-values for all tests where the range profile was rejected as being Pareto Distributed. Recall that the P-value, in simple terms, is the probability of getting a test rejection under the hypothesis the data set is Pareto. Consult [2] for clarification of statistical hypothesis testing. For each range profile an individual Pareto Distribution was fitted for this testing, using Matlab^{TM}. Each statistical test was performed with a test size of 0.05[2]. Table 3.1 also shows the results for both the HH and VV polarisations. In this situation, 14 profiles were rejected out of 966 for the HH channel data, while only 7 were rejected for the VV channel. This corresponds to rejections of 1.5% and 0.72%, which is statistically insignificant. Observe also that the P-values are not extremely small in cases where the range profile is rejected as Pareto. Subsequent re-running of these sequential Kolmogorov-Smirnoff tests, involving different decorrelations, resulted in variations in these results, but on average, the same number of samples were rejected.

[2]The test size is the probability of rejecting the data as Pareto, when in fact it is Pareto Distributed.

Table 3.1: *P-values for testing whether the clutter in run34690_rccal_255 is Pareto. These are the P-values for rejected batches, showing both polarisations.*

Test 1 P-Values	
HH	0.0158 0.0182 0.0350 0.0397 0.0397 0.0238 0.0449
	0.0238 0.0350 0.0350 0.0449 0.0271 0.0309 0.0350
VV	0.0350 0.0449 0.0158 0.0271 0.0182 0.0449 0.0120

Table 3.2: *P-values associated with tests on run34690_rccal_005*

Test 2 P-Values	
HH	0.0120 0.0208 0.0238 0.0309 0.0309 0.0397 0.0449
	0.0104 0.0449 0.0271 0.0449
VV	0.0182 0.0042 0.0104 0.0309 0.0350 0.0208 0.0350
	0.0271 0.0449 0.0238 0.0138 0.0238 0.0067

The second test is for the data set run34690_rccal_005, which contains 922 pulses. Table 3.2 shows the P-values for the rejected samples. In this example, 11 (HH channel) and 13 (VV channel) range profiles were rejected respectively. Relative to the number of pulses, this corresponds to 1.19% and 1.41% rejection rates for the two respective channels, which is small. Again, statistical fluctuations were observed on subsequent test re-runs.

The third sequential tests were performed on the set run34683_rccal_225, which consisted of 821 pulses. Table 3.3 shows both channels had the same rejection rate (0.85%).

Table 3.3: *P-values for goodness of fit tests on run34683_rccal_225*

Test 3 P-Values	
HH	0.0449 0.0208 0.0397 0.0271 0.0350 0.0309 0.0120
VV	0.0238 0.0104 0.0397 0.0238 0.0309 0.0238 0.0350

Table 3.4: *P-values for run34683_rccal_190*

Test 4 P-Values	
HH	0.0449 0.0449 0.0449 0.0271 0.0104 0.0350 0.0449 0.0309 0.0182 0.0350 0.0397 0.0271 0.0397 0.0449 0.0449 0.0077 0.0158
VV	0.0104 0.0397 0.0208 0.0449 0.0138 0.0309 0.0120 0.0397

The final data set considered was run34683_rccal_190, which consists of 832 pulses, and Table 3.4 shows the corresponding P-values for rejected samples. Here, we see that 17 out of 832 (2.04%) range profiles were rejected for the HH-polarised clutter, while 7 (0.84%) were rejected for the VV channel.

3.2 Distributional Fitting: Range Profile Samples

To complement the analysis in the previous subsection, a series of range profiles are analysed graphically. In each case, a Pareto Distribution is fitted with Matlab^{TM} using its inbuilt MLE for the Pareto parameters. The Pareto parameters recorded in the following correspond to those in the density (2.3). Using the results in [16], both K- and KK-Distributions are also fitted to the data for comparison purposes.

The first example is for the range profile corresponding to the 10th pulse in run34683_rccal_225, for the HH channel. The appropriate K-Distribution has scale parameter $c = 40$ and shape parameter $\nu = 4.158$. The KK-Distribution has scale parameters $c_1 = 35$, $c_2 = 114.45$, and the same shape parameter as for the fitted K-Distribution (following the modelling approach outlined in [16]). Also, the mixing parameter was chosen to

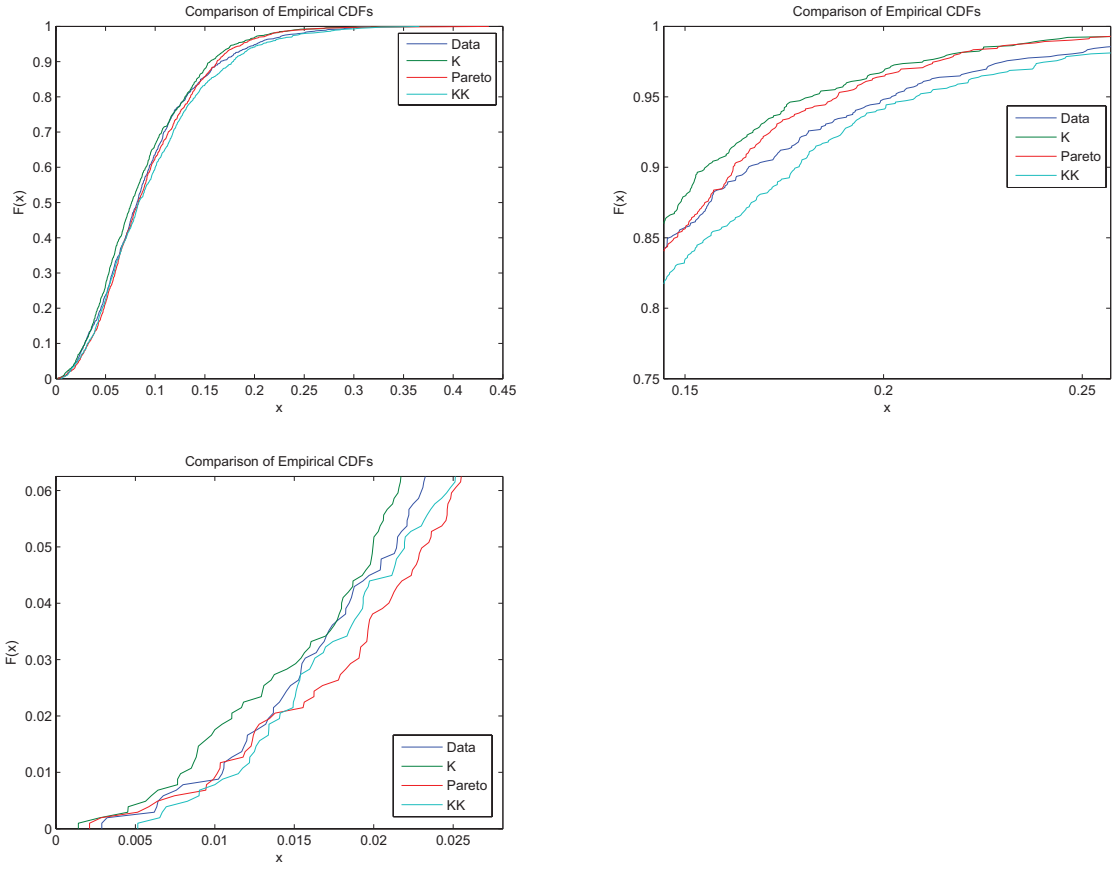

Figure 3.2: *Range profile modelling for run34683_rccal_225 (HH channel), 10th pulse. The figure shows empirical CDF comparisons.*

be $k = 0.01$, as specified in [16]. Matlab^{TM} fitted a Pareto Distribution with parameters $\omega = -0.2418$ and $\lambda = 0.0088$. Figure 3.2 shows plots of the empirical cumulative distribution functions (ECDFs). The first subplot in Figure 3.2 (top left) shows the ECDFs; the top right subplot shows a magnification in the upper distribution region. The latter shows the Pareto is a moderately good fit to the data, but asymptotically like the K-Distribution fit. Here the KK-Distribution is a better fit in the upper tail region. The bottom left subplot shows the lower tail behaviour; we see the Pareto is performing well, but still the KK-Distribution is the best overall fit.

The second range profile considered is for the same data set used previously, but instead we examine the VV polarised clutter. Here the K-Distribution fit has parameters $c = 32$ and $\nu = 9.629$. The results in [16] state that, for this clutter set, the corresponding K-Distributional fit has scale parameters $c_1 = c_2$, and with the same shape parameter as for

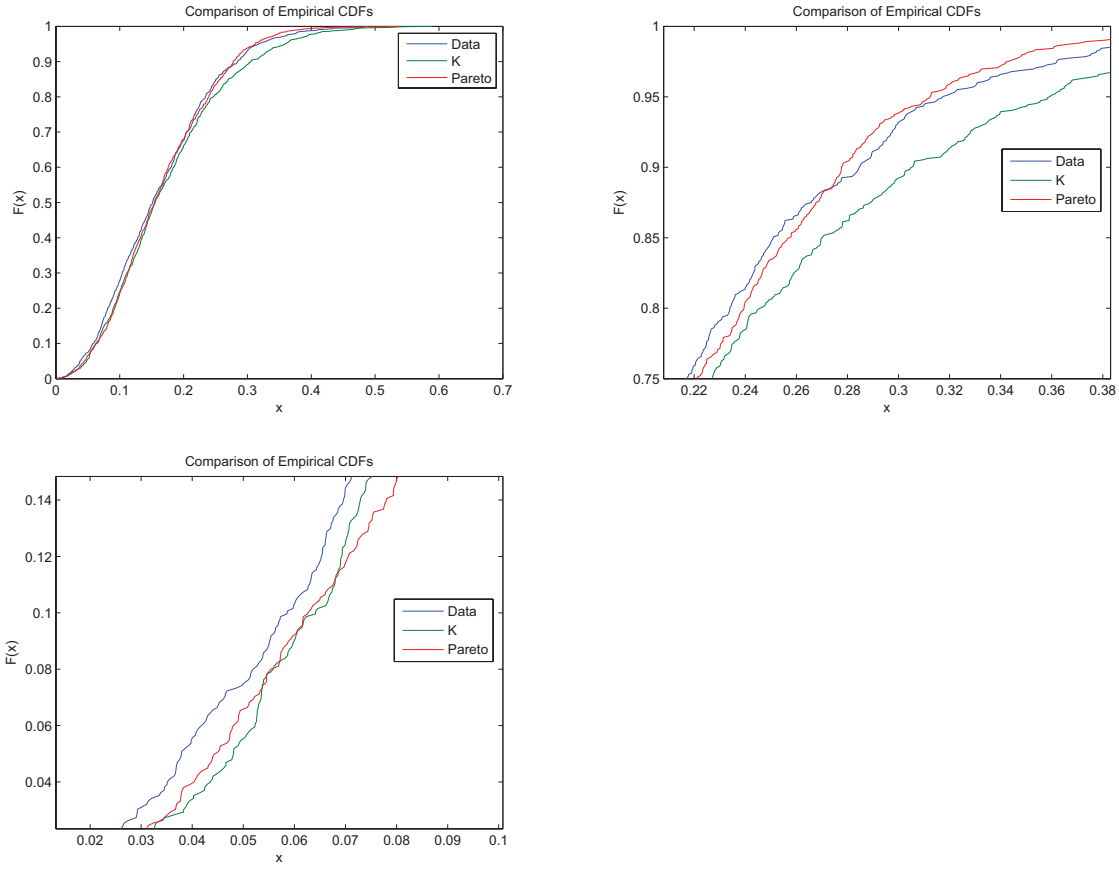

Figure 3.3: *Range profile modelling for run34683_rccal_225 (VV channel), 10th pulse*

the K-Distributional fit. Hence it is exactly the same as the K-Distributional fit. This is not entirely surprising because the KK-Distribution is a model for the spiky clutter found in the HH-polarised channel, while the K-Distribution is still generally a good fit for the VV-polarised channel. Matlab^{TM} fitted a Pareto Distribution with $\omega = -0.040$ and $\lambda = 0.0328$. Figure 3.3 shows the comparison of ECDFs. The upper tail plot (top plot, right) shows the Pareto is better in the upper tail region. It is also good in the lower tail region (see the subplot on bottom left in Figure 3.3), although it is close to the K-Distributional fit.

The third range profile considered is for run34683_rccal_190, and is for a case of HH polarised clutter. Here a K-Distribution with $c = 60$ and $\nu = 4.684$ was fitted. The KK-Distribution had parameters $c_1 = 55, c_2 = 317.35$ and $k = 0.01$. The same shape parameter was used. Matlab^{TM} fitted a Pareto Distribution with $\omega = -0.2605$ and

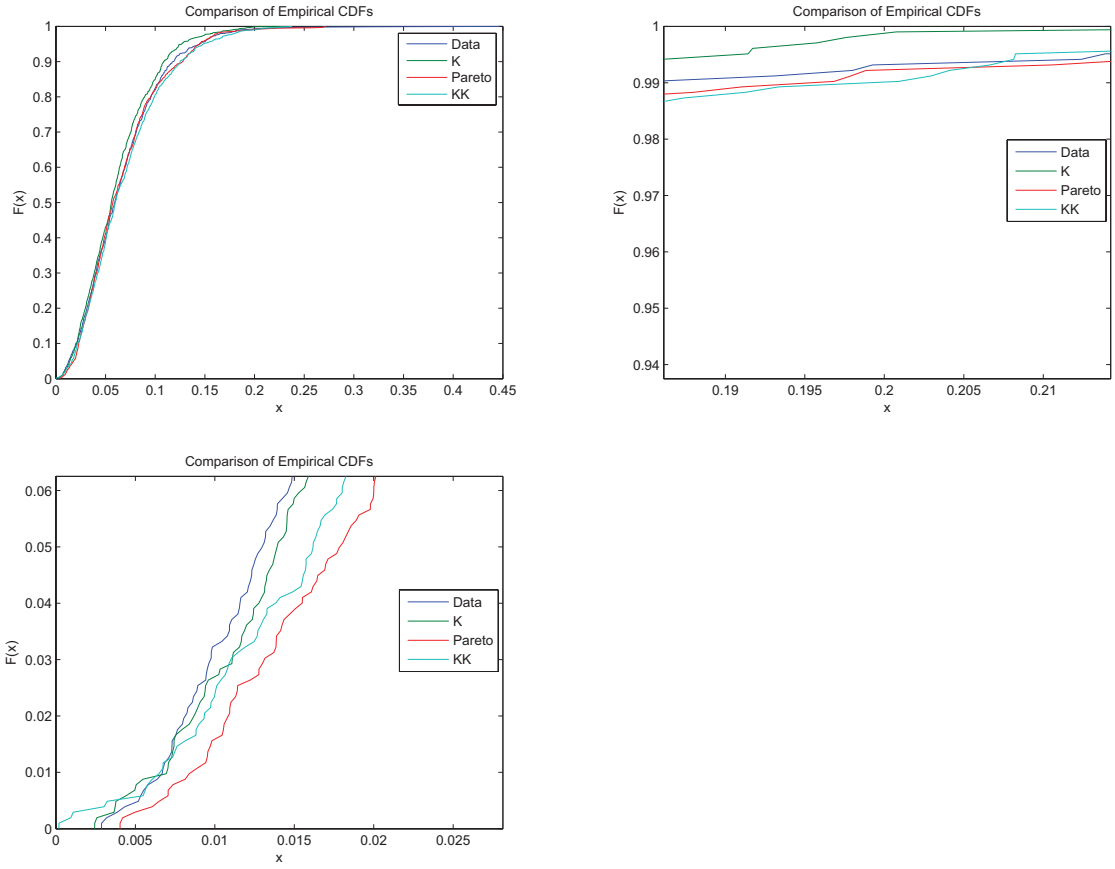

Figure 3.4: *Empirical distribution functions for run34683_rccal_190 (50th pulse, HH polarisation)*

$\lambda = 0.0045$. Figure 3.4 shows the ECDF comparisons. We observe that the Pareto is quite consistent in the upper tail region plot (top right plot in Figure 3.4), while it is not so accurate in the lower tail region (Figure 3.4, bottom left plot).

The final range profile analysed is also for the data set run34683_rccal_190, but for the VV channel. Here the K-Distribution has parameters $c = 50$ and $\nu = 8.556$, while the KK-Distribution has same shape parameter, but used scale parameters $c_1 = 48$ and $c_2 = 129.12$, and mixing parameter $k = 0.01$. The fitted Pareto Distribution has $\omega = -0.0806$ and $\lambda = 0.01441$. Figure 3.5 shows the ECDF plot comparisons. The upper tail region plot shows the Pareto Distribution fits the data rather well (top right plot in Figure 3.5). Also, in the intermediate stage, there is good fitting (bottom right subplot in Figure 3.5). At the lower tail region, there is a fair bit of variation in the fit (bottom left subplot in

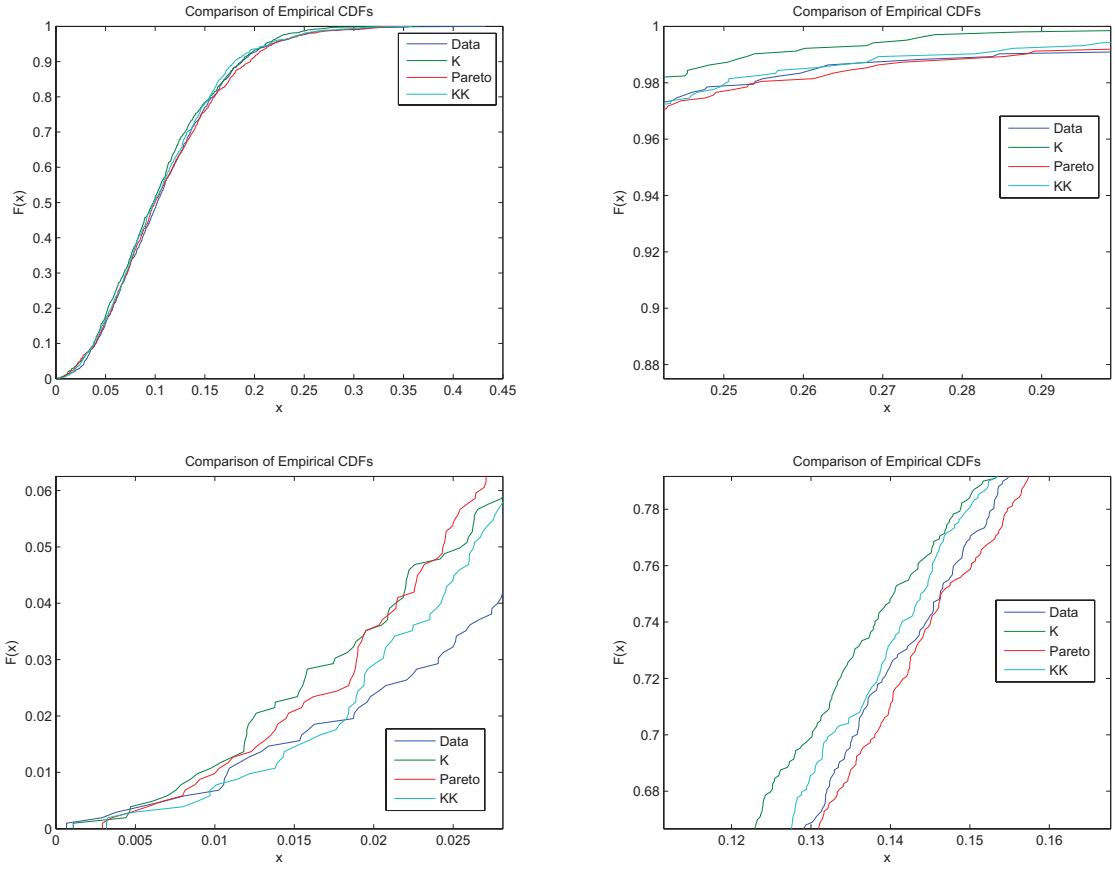

Figure 3.5: *Empirical distribution functions for run34683_rccal_190 (50th pulse, VV polarisation)*

Figure 3.5).

The overall conclusion from the range profile fitting of distributions is that the Pareto model seems reasonable and can perform as well as the K- and KK-Distributional fits. It is important to note that the sample size used for each range profile is 1024, which is insufficient to fully measure the success of a Pareto fitted model.

3.3 Distributional Fitting: Entire Data Samples

This subsection considers the fit of the Pareto distribution to an entire data stream, corresponding to each of the four subsets of data under analysis. Each of the data sets contains azimuth range profiles from 5° to 355°, in steps of 5°. Hence, there will be temporal correlations across adjacent range profiles. No attempt was made to remove this

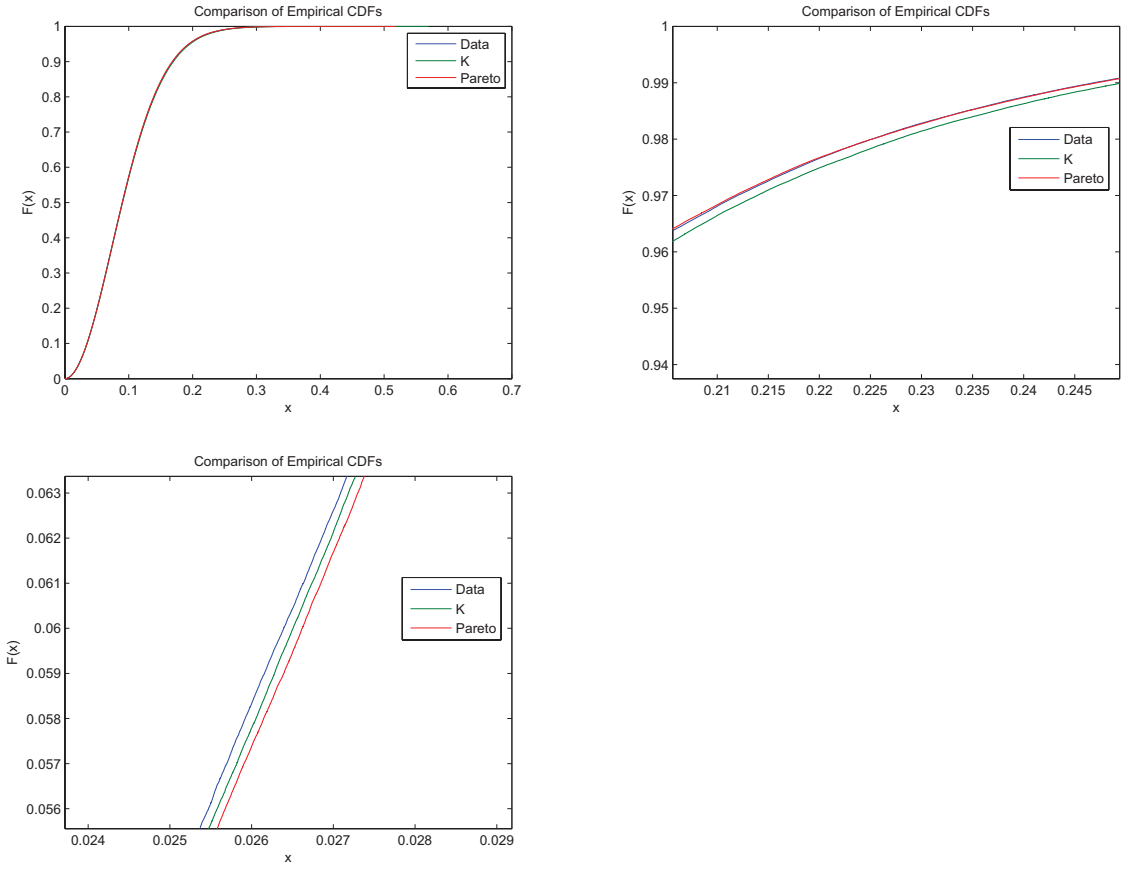

Figure 3.6: *Entire fit of run34690_rccal_255 (VV) data*

in the following, instead the MLE introduced previously for the Pareto parameters were used as an approximation.

The first example is illustrated in Figure 3.6. This has been produced from the VV polarised data in run34690_rccal_255. The K-Distribution has parameters $c = 65$ and $\nu = 13.059$, and the results in [16] state that for a KK-Distributional fit, $c_1 = c_2$. Hence the KK-Distribution fit is exactly the K-Distribution fit in this case. Matlab^{TM} fitted a Pareto Distribution with $\omega = -0.0629$ and $\lambda = 0.0114$. Here the increased sample size shows tighter distributional approximations. The upper tail plot (top right subplot in Figure 3.6) shows the Pareto to perform extremely well; in the lower tail, the K-Distribution has the best performance (bottom left subplot).

The second case considered is for the same data set but for the HH polarised clutter. Here, the K-Distribution has parameters $c = 75$ and $\nu = 2.863$. The KK-Distribution has

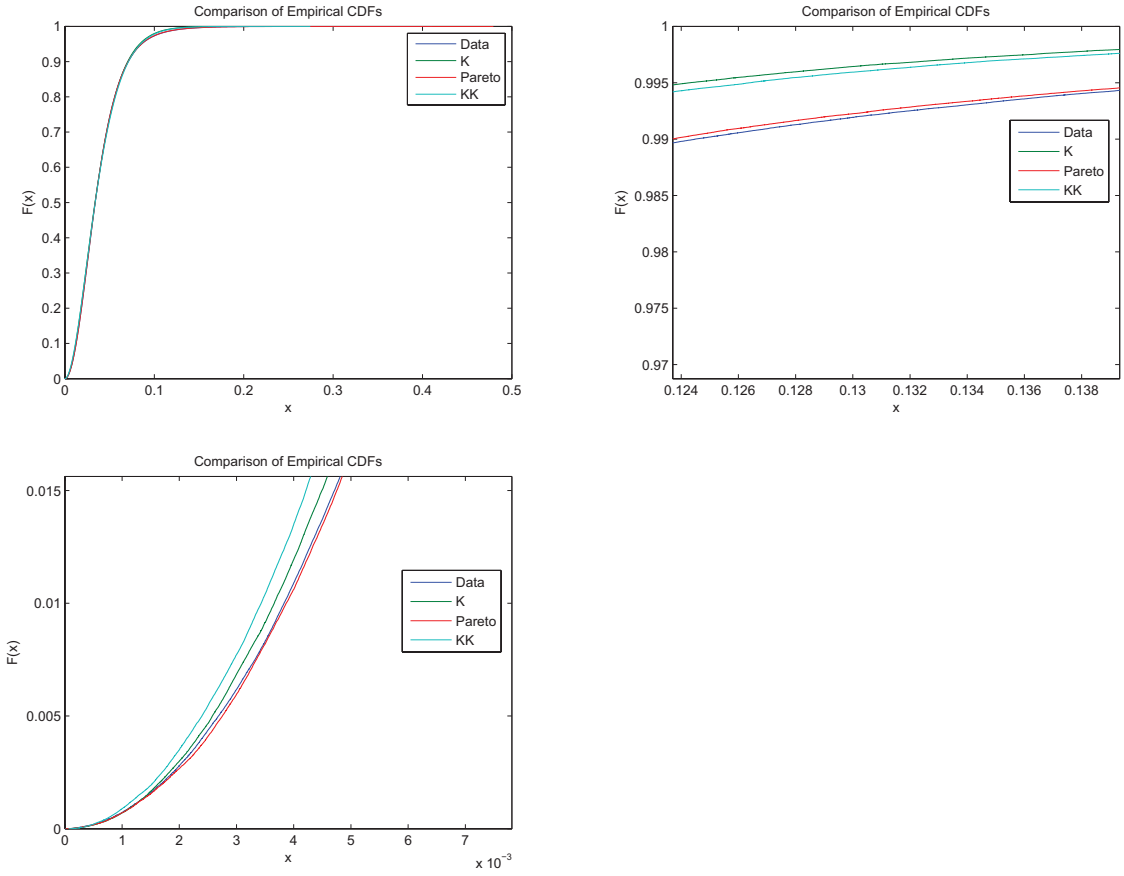

Figure 3.7: *Entire fit of run34690_rccal_255 (HH) data*

the same shape parameter, and scale parameters are $c_1 = 74$ and $c_2 = 356.68$. The mixing parameter is $k = 0.01$. The fitted Pareto has $\omega = -0.3134$ and $\lambda = 0.0015$. The ECDFs are shown in Figure 3.7, which shows the Pareto distribution fits better than the other two models.

Example 3 is illustrated through the ECDF plots in Figure 3.8. This is for the HH polarised clutter in run34683_rccal_190. Here the K-Distribution has $c = 60$ and shape parameter $\nu = 4.684$. The KK-Distribution has $c_1 = 59$ and $c_2 = 340.43$, with same shape parameter. The mixing parameter is again $k = 0.01$. Matlab^{TM} fitted a Pareto Distribution with $\omega = -0.2040$ and $\lambda = 0.0043$. This is another example where the Pareto Distribution provides the best fit.

A final comparison of results is given in Figure 3.9. Here the fitting is for HH-polarised clutter of run34690_rccal_005. The K-Distribution has parameters $c = 145$ and $\nu = 3.2$.

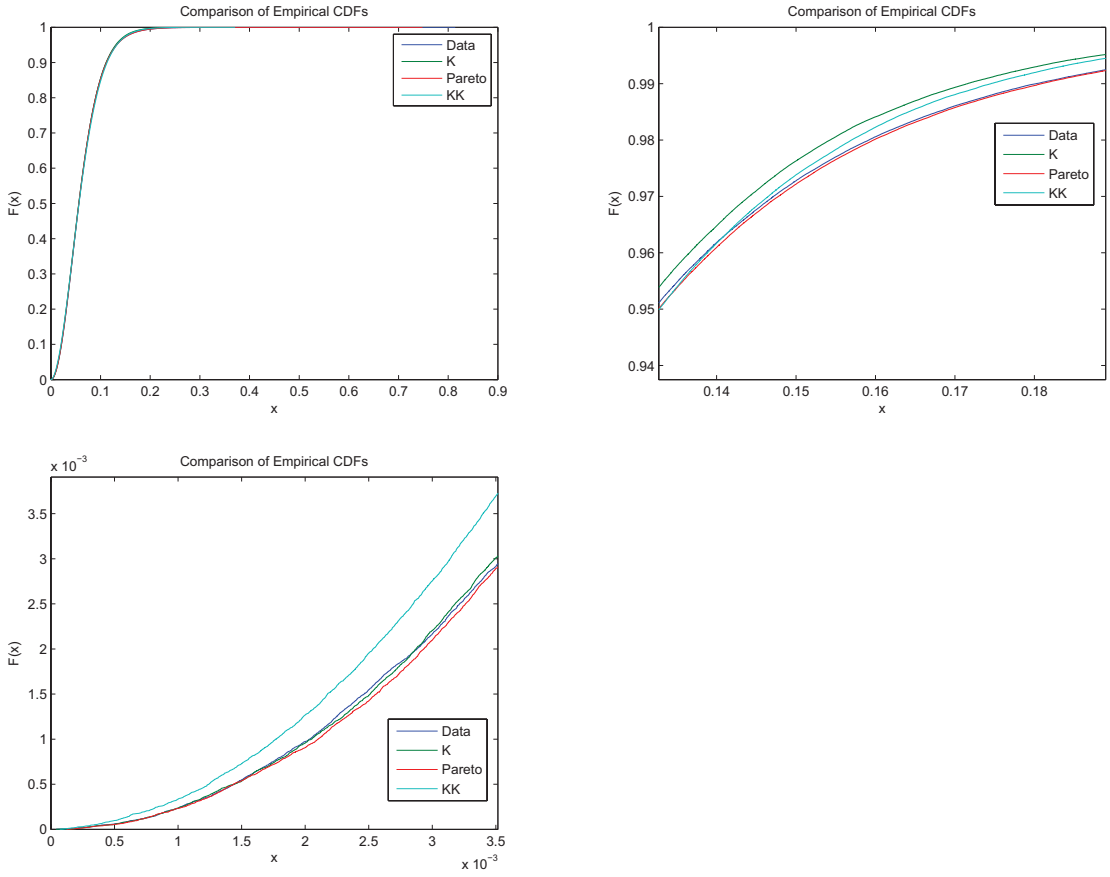

Figure 3.8: *Comparison of ECDFs for run34683_rccal_190 (HH)*

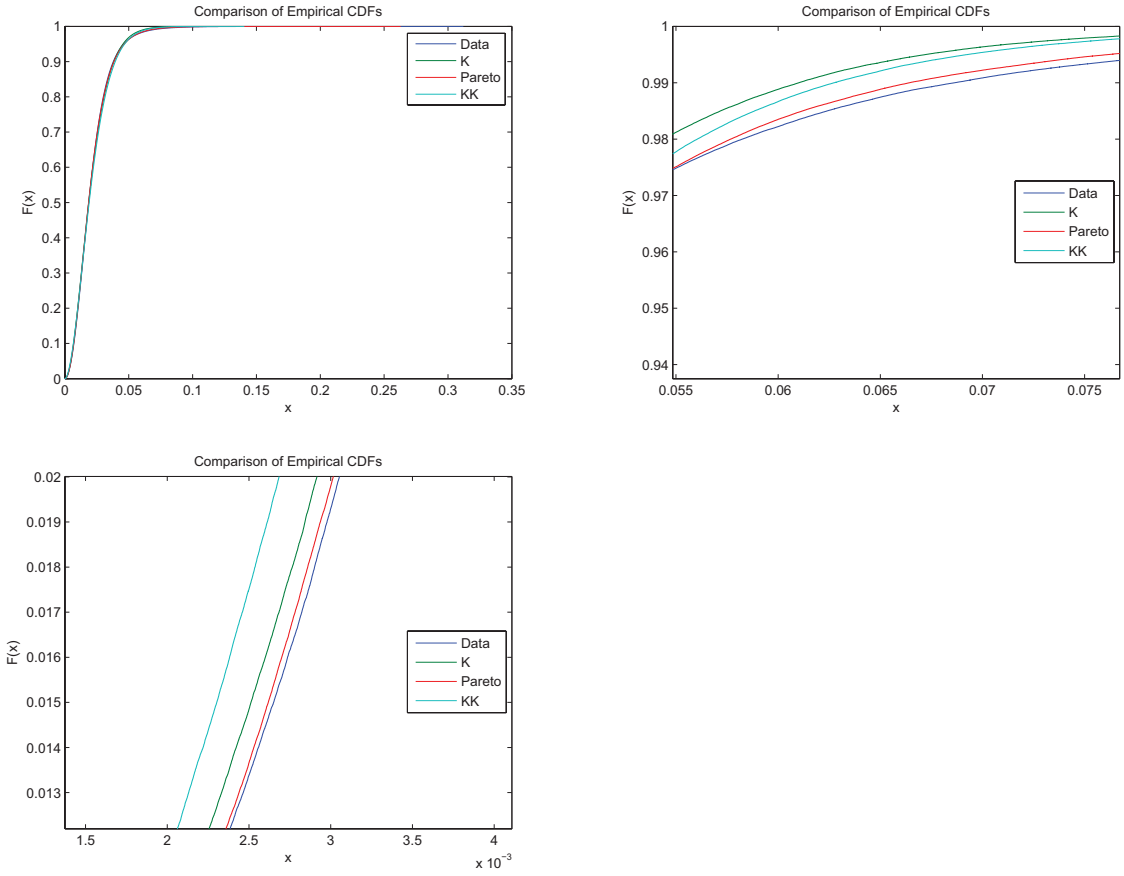

Figure 3.9: *Clutter fitting for HH-Polarised clutter of run34690_rccal_005*

The KK-Distribution has scale parameters $c_1 = 141$ and $c_2 = 877.02$. As before, $k = 0.01$. The fitted Pareto has $\omega = -0.2999$ and $\lambda = 0.0004$. As before, the Pareto distribution provides the best overall fit to the data.

In view of these results, we conclude that the Pareto Distribution can model the high grazing angle clutter quite well, and seems to outperform the K and KK-Distributions in a number of cases.

4 Conclusions

In line with the recent proposal in [13] to use the Pareto Distribution for high resolution radar clutter, its suitability as a model for high grazing angle clutter was considered. It was found that it replicated the reported success in [13]. Range profile fits suffered

because of small sample sizes, but nonetheless, the Pareto Distribution fitted reasonably well. Fitting the Pareto Distribution to an entire clutter set, which encompasses a full 360° scan, showed considerable success. Further tests are needed to validate these results. Overall, it was found the Pareto fitted as well as the K-Distribution, and on larger clutter sets, it outperfomed the KK-Distribution.

Given the relative simplicity of the Pareto clutter model, this work shows that it is possible to base the construction of detectors on this clutter model, which should facilitate the mathematical analysis. It is probably more robust to begin with a simpler clutter model, since suboptimal detectors will have to be introduced subsequently, and it is envisaged these will be simpler to implement under the Pareto clutter model assumption. Current experience has shown the KK-Distribution detectors to be quite difficult to construct [18].

Acknowledgement

Thanks are due to Dr Paul Berry for reviewing this report.

References

1. Arnold, B. C., Pareto Distributions, (International Cooperative Publishing House, Burtonsville, MD, 1983).

2. Beaumont, G. P., Intermediate Mathematical Statistics, (Chapman and Hall, London, 1980).

3. Evans, M., Hastings, N. and Peacock, B., Statistical Distributions, 3rd Edition, (Wiley, New York, 2000).

4. Bruno, G. "Pareto, Vilfredo" The New Palgrave: A Dictionary of Economics, **5**, 799804, 1987.

5. Cirillo, R. The Economics of Vilfredo Pareto, (Frank Cass Publishers, 1978)

6. Aban, I. B., Meerschaert, M. M. and Panorska, A. K., Parameter Estimation for the Truncated Pareto Distribution, Journal of the American Statistical Association, **101**, 270-277, 2006.

7. Asimit, A. V., Furman, E. and Vernic, R., On a Multivariate Pareto Distribution. Insurance: Mathematics and Economics, **46**, 308-316, 2010.

8. Rytgaard, M., Estimation in the Pareto Distribution, ASTIN Bulletin, **20**, 201-216, 1990.

9. Chlebus, E. and Ohri, R., Estimating Parameters of the Pareto Distribution by Means of Zipf's law: Application to Internet Research, IEEE Globecom Proceed., 1039-1043, 2005.

10. Gelb, J. M., Heath, R. E. and Tipple, G. L., Statistics of Distinct Clutter Classes in Midfrequency Active Sonar, IEEE Oceanic Eng. **35** 220-229, 2010.

11. Chakravarthi, P. R. and Ozturk, A., On Determining the Radar Threshold from Experimental Data, IEEE Proceed., 594-598, 1991.

12. Piotrkowski, M., Some Preliminary Experiments with Distribution-Independednt EVT-CFAR based on Recorded Radar Data, IEEE, 2008.

13. Farshchian, M. and Posner, F. L., The Pareto Distribution for Low Grazing Angle and High Resolution X-Band Sea Clutter, IEEE Radar Conf., 789-793, 2010.

14. Stacy, N., Badger, D., Goh, A., Preiss, M. and Williams, M., The DSTO Ingara Airborne X-Band SAR Polarimetric Upgrade: First Results. *IEEE Proceed. Int. Symp. Geo. Remote Sensing (IGARSS)* **7**, 4474-4475, 2003.

15. Crisp, D. J., Stacy, N. J. S. and Goh, A. S., Ingara medium-high incidence angle polarimetric sea clutter measurements and analysis, DSTO-TR-1818, 2006.

16. Dong, Y., Distribution of X-Band High Resolution and High Grazing Angle Sea Clutter. DSTO-RR-0316, 2006.

17. Shorak, G.R. and Wellner, J.A., Empirical Processes with Applications to Statistics, (Wiley, 1986).

18. Weinberg, G. V., Optimal and Suboptimal Radar Detection Of Targets From High Grazing Angles and in High Resolution Sea Clutter, (MSc Thesis, University of Adelaide, 2010).

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. CAVEAT/PRIVACY MARKING

2. TITLE	3. SECURITY CLASSIFICATION
An Investigation of the Pareto Distribution as a Model for High Grazing Angle Clutter	Document (U) Title (U) Abstract (U)
4. AUTHOR	5. CORPORATE AUTHOR
Graham V. Weinberg	Defence Science and Technology Organisation PO Box 1500 Edinburgh, South Australia 5111, Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO–TR–2525	AR 014-947	Technical Report	March 2011

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. No. OF PAGES	12. No. OF REFS
2011/1020762/1	07/040	DGAD	21	18

13. URL OF ELECTRONIC VERSION	14. RELEASE AUTHORITY
`http://www.dsto.defence.gov.au/publications/scientific.php`	Chief, Electronic Warfare and Radar Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SOUTH AUSTRALIA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DSTO RESEARCH LIBRARY THESAURUS

Clutter Modelling	Pareto Distributions
High Grazing Angle Models	Distributional Fitting

19. ABSTRACT

The Pareto Distribution is proposed as a new model for high resolution, high grazing angle radar sea clutter returns. It is assessed using the Ingara radar clutter gathered in 2004. Included is range profile modelling as well as global distributional fitting to clutter streams.